University of Reading

Preferences for mitigation of the negative impacts of the oil and gas industry in the Niger Delta region of Nigeria

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Ukpong, I. G., Balcombe, K. G., Fraser, I. M. and Areal, F. J. (2019) Preferences for mitigation of the negative impacts of the oil and gas industry in the Niger Delta region of Nigeria. Environmental and Resource Economics, 74 (2). pp. 811-843. ISSN 1573-1502 doi: 10.1007/s10640-019-00349-4 Available at https://centaur.reading.ac.uk/83414/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1007/s10640-019-00349-4

Publisher: Springer

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

University of Reading

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Environmental and Resource Economics (2019) 74:811–843
https://doi.org/10.1007/s10640-019-00349-4

Preferences for Mitigation of the Negative Impacts of the Oil and Gas Industry in the Niger Delta Region of Nigeria

I. G. Ukpong[1] · K. G. Balcombe[1] · I. M. Fraser[2] · F. J. Areal[1]

Accepted: 7 May 2019 / Published online: 14 May 2019
© The Author(s) 2019

Abstract
We examine Nigerian preferences for the mitigation of negative impacts associated with oil and gas production using a discrete choice experiment. We analyse the data using a Bayesian 'infinite mixtures' model, which given its flexibility can approximate an array of existing model specifications including the mixed logit and finite mixture specifications. The application of this model to our data suggest multimodality in the marginal willingness to pay distributions associated with mitigation policy characteristics. Individuals are willing to pay for mitigation of negative impacts, but are not necessarily willing to trade-off very large increases in unemployment or poverty to achieve these benefits.

Keywords Discrete choice experiment · Bayesian infinite mixture · Dirichlet process · Nigeria

JEL Classification O13 · Q34 · C11

1 Introduction

Nigeria is a large producer and exporter of crude oil, historically the largest in Africa (BP 2015). The oil and gas (O&G) industry contributes over 60% of the country's national budgetary revenue, 15% of GDP and 95% of foreign exchange earnings through crude oil exports (NBS 2012). Crude oil was first discovered in Nigeria in 1956 in Bayelsa State in the core Niger Delta region by Shell Darcy, now known as the Shell Petroleum Development Company of Nigeria. This discovery had a significant impact on the economy of the region which traditionally has been dependent on the natural environment, with farming, forestry and fishing being the major economic activities. However, the region has

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10640-019-00349-4) contains supplementary material, which is available to authorized users.

✉ F. J. Areal
f.j.areal@reading.ac.uk

[1] School of Agriculture, Policy and Development, University of Reading, Whiteknights, PO Box 237, Reading RG6 6AR, UK

[2] University of Kent, Canterbury, Kent CT2 7NP, UK

experienced numerous negative externalities as a result of the activities of the O&G industry including soil and water pollution which have lead to habitat and biodiversity loss estimated to be as high as US$11,000 per household (Adekola et al. 2015). These externalities have impacted the livelihoods of local communities through losses in production from fishing, farming and forestry (Idemudia 2010) leading to poverty, unemployment and outward migration. There is also conflict with attacks on oil transport pipelines, violent protests, attacks on oil installations and kidnapping of oil workers, with insecurity in the region an international concern (Omofonmwan and Odia 2009). It is estimated that between 2000 and 2010 the Nigerian Federal government lost in excess of 170 billion Naria due to such conflicts (Anifowose et al. 2012).[1]

In response, the Nigerian Federal government has enacted various environmental laws and attempted to regulate the operations of the O&G sector so as to reduce the negative consequences of their activities (Idemudia 2010). In addition, the O&G sector has also adopted measures to minimise negative environmental externalities and socio-economic problems associated with O&G production, including the payment of compensation to affected individuals and communities, and recognising corporate responsibilities (Idemudia 2014). However, despite these efforts, the Niger Delta region remains vulnerable to O&G related environmental problems and associated economic hardship.

Using the Nigerian O&G sector as a case study, we make two contributions to the valuation literature.

First, we examine the potential opportunity costs and benefits associated with O&G activity with a view to highlighting potential priorities in order to inform policy formulation and regulation. Although the activities of the O&G sector impact local communities regarding the environment there are associated benefits in terms of employment and poverty reduction that need to be taken account of in developing appropriate policy responses. Currently, it is unclear if benefits for maintaining employment and minimising poverty are greater than or less than the costs associated with reducing environmental externalities. To start to address these important policy trade-offs, we employed a Discrete Choice Experiment (DCE) in the Niger Delta region of Nigeria. To date, virtually no stated preference research on this issue or any other environmental issue has been conducted in Nigeria, with the only exception being Urama and Hodge (2006) who undertook a contingent valuation study to examine farmers' attitudes and willingness to pay (WTP) for a river basin management scheme.[2]

In undertaking this DCE, the set of attributes employed take account of societal trade-offs such as job creation (e.g. Colombo et al. 2005; Longo et al. 2008; Kosenius and Ollikainen 2013). Given the importance of the O&G sector, we considered it important to ask respondents not only to think about attributes that have private costs and benefits, but also attributes that have strong social consequences, including unemployment and poverty. In doing this, we recognise that the role of altruism and its place in valuation has raised a series of issues (e.g. Haab et al. 2013; Zhang et al. 2013). A strand of the literature has investigated the appropriateness of allowing respondents to consider welfare effects on the rest of society, distinguished between types of altruism, and even argued that some forms of altruism have no place in valuation (e.g. McConnell 1998; Flores 2002; Jacobsson et al. 2007). Our position is that it is legitimate (and desirable) to include attributes with social

[1] The current exchange rate is £1 = N304.48.

[2] The absence of non-market monetary values for Nigeria and the Niger Delta in particular is noted by Adekola et al. (2015) in relation to global efforts to value ecosystems such as wetlands.

benefits within this DCE. Indeed, most stated preference studies for non-market environmental goods and services require trade-offs between attributes that respondents will recognise to have an immediate or direct on themselves but also an impact on others. To the extent that respondents value outcomes that benefit or hurt others, this will be reflected in their marginal utilities (albeit it may not be possible to decompose the sources of utility). On a theoretical level this creates difficulties in the decomposition between what constitutes private or social preferences, but it does not ultimately place restrictions on peoples ability to value changes in attributes. Putting theoretical objections aside, excluding attributes with social consequences from an experiment requires the assumption that respondents either do not care about these attributes, or can and do hold these attributes constant when considering changes in the attributes that are included. In many cases, this will be in sharp conflict with what they would expect to happen in reality. Consequently, respondents will tend to factor this in when making their choices, and attempts to enforce unrealistic *ceteris paribus* assumptions are likely to induce dissonance in the minds of respondents when making their choice.

Our second contribution is to the bring to the attention of DCE researchers an alternative way to model respondent preference heterogeneity. We employ a Bayesian 'infinite mixtures' Logit (BIML) to estimate the preference parameters of respondents. To date the majority of the DCE applications in the literature model respondent heterogeneity of preferences by using what has been termed the 'Mixed Logit' (ML) (Hensher and Greene 2003; Fiebig et al. 2010), also referred to, within Bayesian circles, as the Hierachical Bayesian Logit (HBL) (e.g., Balcombe et al. 2016). This approach models preference parameters as continuous preference distributions (e.g., normal or log normal). Alternatively, researchers employ latent class models (LCM) using Classical terminology (e.g., Greene and Hensher 2003) which in a Bayesian setting is implemented using Finite Mixtures (FM). However, there are limitations with both approaches. For example, the ML (HBL) approach struggles to cope with circumstances where respondents fall into distinct groupings or the appropriate choice of distribution with which to model preference heterogeneity. At the same time, the key limitation of the LCM (FM) model is that it struggles to identify the correct number of mixtures (i.e., classes) and may be very poor at characterising cases when there are continuous preference distributions. While on one hand researchers have a multitude of model performance criteria available to them, they often conflict and very little is in fact known about how well these criteria can discriminate between these models (in finite samples). There is therefore little consensus as to precisely which criteria should be employed for model selection, nor consensus about which model should be seen as the leading candidate model.

The BIML potentially provides the means by which to overcome the limitations of both approaches by introducing a greater degree of flexibility with regard to capturing preference heterogeneity. Unlike the HBL, it is not necessary ex-ante to select random parameter distributions because the BIML estimates these non-parametrically. In addition, the use of the term 'infinite' indicates that this model does not require the number of mixtures to be set ex-ante (nor ex-post). The BIML is only 'infinite' in the sense that there is no upper bound on the number of mixtures allowed given infinite data. While a finite sample implies that the number of sample mixtures will be finite, the number may be much larger than typically observed in models reported in the literature. But, just as importantly, the BIML does not seek to select a number of mixtures upon which parameter estimation and inference becomes contingent, but instead treats the number of mixtures as a parameter to be estimated and the uncertainty about its value is embodied in all parameters of interest.

The flexibility afforded by the BIML links to a recent strand of the literature following Train (2008, 2016) who introduced and developed the logit-mixed logit (LML) model specification . The LML has gained traction in the DCE literature because of the flexibility associated with how random parameter distributions can be modelled (e.g. Bazzani et al. 2018; Bansal et al. 2018; Caputo et al. 2018; Bhat and Lavieri 2018). The appeal of the LML is that it reduces the need for researchers to make parametric assumptions that may not be supported by the underlying data generating process. However, as is evident from the applications and extensions of the LML there remain a range of issues when employing this approach that warrant more attention.

The BIML provides an alternative Bayesian but related approach to the LML by providing greater flexibility in modelling DCE data. Importantly, in this paper we do not claim that the BIML is necessarily the best approach to model estimation. If preferences accurately conform to the underlying assumptions of the HBL model or alternatively the LCM with a small number of mixtures, then the use of one or the other of these models will be optimal. However, what we argue here, is that the BIML is a flexible approach that works well where there are small number of classes, or in circumstances where there are a very large number of classes or even continuous distributions. It will not be our aim here to show that the BIML outperforms either the HBL or LCM model. We do, however, provide in "Appendix 1", a simulation study that illustrates how the BIML performs when data generating processes are varied and would by design favour other model specifications. In addition, we compare our BIML results with those from an HBL and FM models for our data, and in doing so illustrate to DCE researchers that this approach is worth further investigation.

We proceed by outlining the econometric specification in Sect. 2. In Sect. 3, we discuss the DCE design, survey, sampling procedure and summarise the sample. Then in Sect. 4, we discuss our choice of model priors, explaining how they impact model performance. Next, we present the results of the BIML along with a discussion in relation to the DCE and in Sect. 6, we conclude.

2 The Econometric Specification

2.1 The Dirichlet Process and Discrete Choice

Within the BIML, individual preference parameters are formed from a mixture of distributions. The estimates for a given individual are constructed as a weighted sum of the distributions that compose the mixture, where the weights are the probabilities that a person belongs to a particular 'class'. The BIML derives from a Dirichlet Process (DP) prior on the number of mixtures.[3] The nature of the DP is sometimes described either in terms of the 'Polya Urn' or 'Chinese restaurant process'. These analogies usefully convey the way in which individuals are allocated to groups, whereby the larger the population of a group, the larger the probability that a given individual will be allocated (or reallocated) to that group. In effect, the DP acts as a prior towards shrinking the number of classes (mixtures), unless a larger number of classes (mixtures) is supported by the data.

[3] Models using DP priors have been given alternative names including Mixture of Dirichlet Process (or MDP) models.

A general description of DP and its relationship with the Dirichlet distribution are described in Frigyik et al. (2010) with other useful discussions in McAuliffe et al. (2006) and Walker et al. (2007). In addition, there have been earlier applications of the DP within the economics literature; see Burda et al. (2008), Daziano (2013), Jin and Maheu (2016) and Bauwens et al. (2015). Here we give an outline of this approach in the context of a discrete choice framework.

As standard for a DCE, the utility U_{ijs} of the jth respondent for the ith option in the sth choice set (i.e. the set of options offered to respondent j) can be expressed as follows:

$$\begin{aligned} U_{ijs} &= u\left(x_{ijs}, \beta_j\right) + e_{ijs} \\ &\quad \text{for} \\ &\quad i = 1, \ldots, N \\ j &= 1, \ldots, J \\ s &= 1, \ldots, S \end{aligned} \tag{1}$$

where x_{ijs} is a $(K \times 1)$ vector of known attributes, and β_j is a vector of parameters characterising individual preferences. Commonly, it is assumed that $u\left(x_{ijs}, \beta_j\right) = x'_{ijs}\beta_j$, but this need not be the case. The assumption of a Gumbel (extreme value) error e_{ijs} that is independent across, i, j and s implies that the probability of choosing option i (i.e. yielding maximum utility) for the jth person from the sth choice set is:

$$p_{ijs}\left(\beta_j\right) = \frac{e^{u\left(x_{ijs}, \beta_j\right)}}{\sum_i e^{u\left(x_{ijs}, \beta_j\right)}} \tag{2}$$

Within this general model structure preference heterogeneity can be treated in a number of ways. If employing the HBL approach researchers commonly assume that $\beta_j \sim N(\mu, \Omega)$ (i.e., a prior distribution is assigned to the latent parameters β_j) such that the closer the posterior distribution is to a normal distribution, the better a model performs. In contrast, with the FM formulation, β_j are assumed to be a finite set of classes. It then follows that there are a finite set $\left\{\beta_m\right\}_{m=1}^{M}$ of vectors such that for each individual there is a class $c_j = m$ such that $\beta_j = \beta_{c_j}$.

The model we explore here does not differ from these models in terms of the way that utility is treated as a linear combination of attributes with Gumbel error. It differs only in the way that it treats the heterogeneity in β_j. As is standard across all of these approaches, we define:

$$y_{ijs} = 1 \text{ if } U_{ijs} = Max\left(U_{1js}, \ldots, U_{N,j,s}\right) \tag{3}$$

and assume that this represents the choices made by individuals. However, when estimating the BIML, each individual in the sample is allocated to a class (mixture) according to a probability measure of belonging to each class at each iteration of the estimation process. The parameters which characterise each grouping are then drawn contingent on that grouping. All respondents are then reallocated membership of a class in the next iteration of estimation and the process continues iteratively. The main difference between the BIML and the FM is that any individual can be allocated to new classes which previously did not exist, and classes which hold no membership are 'forgotten'. Therefore, the number of classes with a non-zero membership can increase or decrease throughout estimation where the number of classes is strictly limited to the number of individuals in the sample.

To show the main difference between the BIML and FM, we first introduce a Bayesian FM logit model with M classes using the definitions introduced:

$$\begin{aligned} y_{ijs}|c_j, \{\beta_m\} &\sim p_{ijs}\left(\beta_{c_j}\right) \\ c_j|\{\pi_m\}, C &\sim Discrete\left(\pi_1, \ldots, \pi_M\right) \\ \beta_m &\sim G_0 \\ \left(\pi_1, \ldots, \pi_M\right) &\sim Dirichlet\left(\frac{\alpha}{M}, \ldots, \frac{\alpha}{M}\right) \end{aligned} \tag{4}$$

where π_m are the probabilities of falling into class m, G_0 is some prior distribution for the parameters (frequently normal) and α is a hyperparameter, that determines the 'concentration' of classes. The main limitation of this model is that it requires the pre-selection of M, the number of classes. However, within the Bayesian literature, a way to determine the number of classes endogenously within the estimation process, is by using a DP model (as distinct to a Dirichlet distribution which is finite).

The DP model is a generalisation of the FM model such that the number of potential classes can be infinite (i.e. $M \to \infty$) in principle. In practice, there will always be a finite number of classes where the number of classes is generated within a Monte Carlo Markov Chain (MCMC) algorithm (this number will change stochastically). The DP model with a Gumbel distribution for the error and for individual $j's$ responses $\left\{y_{ijs}\right\}_{i,s}$ denoted Y_j is:

$$\begin{aligned} Y_j|\beta_{c_j} &\sim f\left(Y_j|\beta_{c_j}\right) = \prod_{s=1}^{S}\prod_{i=1}^{N}\left(p_{ijs}\left(\beta_{c_j}\right)\right)^{y_{ijs}} \\ \beta_{c_j}|G &\sim G \\ G &\sim DP\left(G_0, \alpha\right) \\ \alpha &\sim Gamma(z, v) \end{aligned} \tag{5}$$

where $DP\left(G_0, \alpha\right)$ represents a DP with base distribution G_0 and concentration parameter α. While G_0 can be continuous, G is a discrete distribution. Under the DP, the prior probability that an individual falls within a particular class becomes conditional on the number of members already falling within that class. Econometrically, increases in α will act to increase the number of classes. In principle, α can be set to a particular value for a given data set. In our model specification it is estimated using the hierachical priors following Escobar and West (1995).

2.2 BIML Estimation

The approach we take to model estimation is to employ an MCMC algorithm outlined by Neal (2000).[4] Specifically, we implement a Metropolis Hastings (MH) algorithm with additional steps in the algorithm allocating individuals to classes. It is during these steps in

[4] Other authors such as Burda et al. (2008) have also used the algorithms described by Neal (2000).

the algorithm that a class can be introduced or eliminated, and this process will continue until the model converges.

Our algorithm works as follows. At a given point in the MCMC we define $\{m\}_{m=1}^{M}$ as the set of non-empty classes, that is classes with at least one individual allocated to each class. We also have $\{\beta_m\}_{m=1}^{M}$ and $\{c_j\}_{j=1}^{J}$ as defined already. Next, we define n_m to be the number of individuals allocated to class m and $n_{-j,m}$ the number of individuals allocated class m having excluded the jth individual. Following Neal (2000), we then define the proposed distribution for new classes following an iteration of the MCMC algorithm for each individual as follows:

$$\begin{aligned} \Pr ob\left(c_j^* = m^* | m^* \in \{m\}_{m=1}^{M}\right) &= \frac{n_{-j,m^*}}{J-1+\alpha} \\ \Pr ob\left(c_j^* = m^* | m^* \notin \{m\}_{m=1}^{M}\right) &= \frac{\alpha}{J-1+\alpha} \end{aligned} \tag{6}$$

In Eq. (6) the top probability is that which is proposed such that at the next iteration the jth individual is allocated to an existing class. In contrast, the lower probability in Eq. (6) is that which applies if individual j is proposed to be in a class that does not currently exist. A summary of the the algorithm used for updating classes is as follows:

Step 1 for $j = 1, \ldots, J$ draw candidate $c_j^* = m^*$ from the proposal distribution shown in Eq. (6);
Step 2 if c_j^* is not from an existing class then draw β^* from G_0, otherwise $\beta^* = \beta_{m^*}$;
Step 3 accept $c_j = c_j^*$ with probability [with $f\left(Y_j|\beta_{c_j^*}\right)$ defined in (5)]

$$u = \max\left(1, \frac{f\left(Y_j|\beta_{c_j^*}\right)G_0\left(\beta_{c_j^*}\right)}{f\left(Y_j|\beta_{c_j}\right)G_0\left(\beta_{c_j}\right)}\right); \text{ and}$$

Step 4 redefine $\{m\}_{m=1}^{M}$, $\{\beta_m\}_{m=1}^{M}$, and $\{c_j\}_{j=1}^{J}$ if necessary (that is add new non-empty classes and eliminate empty ones).

This algorithm sits within a MH algorithm for which values of $\{\beta_m\}_{m=1}^{M}$ are updated conditional on the existing number of classes. We initiate this process at an arbitrary starting point and then it is repeated for $t = 1, 2, \ldots, T + T_0$, with the first T_0 points being the 'burn in' phase.

At each iteration t (that is after completing the assignment of individuals to classes and the draw of the parameters associated with those parameters) there is a draw of α from its posterior distribution which is a function of the hyper parameters (z and v) in the prior for α along with its current values, the number of classes (M) and number of observations which in this case are the number of respondents (J). How to draw the posterior distribution for α is outlined in Escobar and West (1995), whereby the following steps need to be taken[5]:

[5] For details behind these derivations see Escobar and West, in particular page 585.

Step 1 Draw $\eta \sim Beta(\alpha + 1, J)$ and calculate $b = (z + M - 1)/(n(v - ln(\eta)))$, and $\pi = \frac{b}{1+b}$; and
Step 2 Draw $\alpha = \pi y_1 + (1 - \pi)y_2$ where $y_1 \sim Gamma(z + M, (v - \ln(\eta))$ and $y_2 \sim Gamma(z + M - 1, v - \ln(\eta))$.

At each iteration t, we can estimate the parameters for each individual by constructing the posterior probabilities for class membership under a uniform prior probability that an individual belongs to any one of the classes. Next we can define $\omega_{j,m}^{(t)}$ as the posterior probability that individual j is in class m at iteration t, such that we can then take the estimate of their parameters to be $\beta_j^{(t)} = \sum_{m=1}^{M(t)} \omega_{j,m}^{(t)} \beta_m^{(t)}$, and these are then recorded for all individuals for $t = T_0 + 1, \ldots, T + T_0$. Finally, our parameter estimates are derived for each individual and are constructed from the mean or median for this distribution.

2.3 BIML and the Number of Classes

With the BIML the number of classes (or mixtures) needs to be viewed as a random parameter within a meta model. As such, we argue that it is reasonably sensible to consider this a parameter that should be estimated. However, researchers currently do not have a reliable way to estimate this parameter using existing methods. This is because:

1. there are either large numbers of classes to be estimated; and/or
2. there classes with overlapping high density regions; and/or
3. if the parameters in the models are better modelled as a continuous distributions rather than a discrete mixture.

Under "idealised" circumstances one can show that a FML or HBL model will do well provided that a researcher possesses knowledge of the true model, generates large quantities of data from that model and then uses the correct model to estimate the data. However, this ignores the central issue; in applied research, we do not know the correct number of classes a priori.

The classical way of approaching the class/mixture dimension issue [until Train (2016)] is most commonly to try and make a singular choice over the number of classes using information criteria. In some circumstances information criteria will do well, and in others they will do badly, depending on the degree of problems associated with the three points noted above. We believe the obvious way around this problem is not to view the central problem as being to select the correct or true number of classes, but to adopt an approach that recognises that the number of classes is inherently uncertain. Subsequently, we can ex-post take a mean or a median as a measure of the number of classes, but unlike the finite approach we do not assume necessarily that there is an optimal number of classes. Furthermore, if the underlying parameter distribution is multivariate normal, there is in one sense no "correct" number of classes. Therefore, what must be appreciated is that the DP prior approach is a way not only to map the class dimension but to place a distribution on the number of classes. Finally, we also observe that the classical instinct would perhaps be to go further and insist that estimation should then be restricted to the modal number of classes. However, the Bayesian response would be that this is simply restricting inference about parameters to a subspace of the posterior without any particularly good reason. For this reason the parameter estimates that are reported are a weighted average of the modal number of classes for the model once it has converged.

3 DCE Survey Design and Data Collection

3.1 The Study Area

The Niger Delta region consists of nine states yielding the bulk of O&G produced in Nigeria. Within the Niger Delta, we focussed on three oil producing states: Akwa Ibom, Bayelsa and Rivers. These states are varied in terms of their economic situation as captured by key economic measures reported for 2011/12. Akwa Ibom is ranked 8*th* out of 36 states in terms of GDP, it is the largest state by O&G production, unemployment stands at 18.4%, and it has a population of 4,625,100. Bayelsa is ranked 26*th* out of 36 states in terms of GDP, unemployment stands at 23.9%, and it has a population of 1,970,500. Finally, Rivers is ranked 2*nd* out of 36 states in terms of GDP mainly due to the presence of Port Harcourt and its importance to the O&G sector, unemployment stands at 25.5%, and it has a population of 6,162,100.

3.2 DCE Development

3.2.1 Focus Groups and In Depth Interviews

The development of the DCE began with a series of semi-structured focus groups which were qualitatively analysed in order to understand public perceptions of the impacts of the O&G industry in the study area. The purpose of this analysis was in part to identify and describe the main concerns regarding the O&G industry by different groups in the region, but also to inform the selection of attributes for the DCE. We undertook three focus groups between March and April, 2013, with one each of the three states selected. In total 31 participants (20 male and 11 female) with different backgrounds from communities and Local Government Areas (LGAs) across the three states participated in the sessions.

In addition to the focus groups, in depth interviews were conducted with 10 environmentalists and 10 O&G industry workers.

Both the focus groups and the in depth interviews began with a teaser question:

"If you have 1 million Naira as a gift from a friend, what comes into your mind first?"

The purpose of the teaser question was simply to create an interactive atmosphere before the main discussion started. The first question used to structure the discussion of the focus groups and in depth interviews was as follows:

"How does the oil and gas industry positively impact on people's livelihood in your area?"

To understand the responses provided, we employed content analysis so as to identify themes which could be linked to outside variables such as the role of participant (Ritchie and Lewis 2011). Two main themes regarding participants' concerns about the operations of the O&G industry were identified. These were:

1. the socio-economic impacts of the gas and oil industry; and
2. the environmental impacts associated to the O&G industry activities.

A number of sub-themes were also identified and concerns did vary according to respondent type. However, the most important theme that emerged was that the majority of respondents view the activities of the O&G industry to be in competition with the local communities regarding the environment. Thus, although we might consider the benefits

associated with O&G industry activity in terms of employment and poverty reduction to be important, this cannot be taken for granted from the point of view of public perception. In the words of one focus group participant;

"oil companies have continuously destroyed the major sources of income of people, mainly land, forest and rivers which has brought poverty to the people, and have affected the living standards of people in the rural communities".

This type of statement which emphasised the economic and environmental impact of the O&G industry was frequently repeated. Thus, there is an understanding of the trade-offs involved with having O&G sector operating in the Niger Delta.

Unsurprisingly, environmentalists' major concerns regarding the O&G industry relate to the environmental impact. This group was less concerned with economic losses for the O&G industry even if they are the result of vandalism and sabotage. Nor did environmentalists emphasise the concerns of the focus group members in relation to the O&G industry and the negative impacts on rural livelihoods. Rather, the environmental impacts were viewed by environmentalists to be independent of the impact on peoples livelihoods. In contrast, the O&G workers did not acknowledge the problems facing the wider community associated with the extraction of O&G. While this does not indicate that they are unaware of public concerns, they were very much secondary to the problems they face as a result of working in the industry, or their personal and economic security. Also, the focus group participants and environmentalists seemed dispassionate about issues of most concern to O&G workers such as kidnapping, militancy, vandalism and oil theft.

Overall the varying concerns of the groups consulted, as well as their different attitudes towards the O&G industry and associated impacts, brought into focus the important trade-offs that needed to be considered when devising the DCE.

Finally, another important discussion that took place during the focus group research and the in-depth interviews was whether or not we should employ a willingness-to-accept (WTA) question in the DCE or a WTP question. If we assume that environmental policy should be based upon the implementation of the polluter pays principle then a WTA question based on potential compensation for damages inflicted would be the appropriate way to frame the DCE question. However, it transpired during the focus groups that in general respondents did not consider it unreasonable that they could or should pay for mitigating the negative impacts of O&G. One way to explain this result is that because the public have been subject to environmental degradation for so long, the O&G industry has obtained an implicit property right to pollute because the reference point of society is based upon the historical existence of polluting activities. Thus, we framed the DCE question in terms of WTP.

3.2.2 DCE Attributes

Our DCE is designed primarily to estimate respondents WTP for alternative forms of mitigation activity targeted towards reducing the negative impacts of the O&G sector in the Niger Delta region.

Before we arrived at the final form of the DCE a trial version of the survey was implemented as a pilot study. This was done in two phases in the study area. The first phase of the pilot survey was carried out using convenience sampling. A total of 58 questionnaires were completed and used for the preliminary analysis. After adjustment, a second pilot survey was carried out to pre-test the refined questionnaire. This involved 15 participants, with five participants from each of the three states.

Table 1 Choice experiment attributes

Attribute (units)	Name	Base level	Levels
Tax (Monthly increase in Nigerian Naira)	TAX	0	0, 100, 200, 300, 400, 500
Land occupied by oil and gas pipelines (km)	LANDOCC	4500 km	0, 10, 20, 30, 40, 50[a]
Unemployment rate in Nigeria (%)	UNEMP	20%	20, 25, 30, 35
Number of oil spills	SPILL	320	0, 10, 20, 30, 40, 50, 60*
Amount of gas flared (Billion cubic feet/day)	FLARE	2.5	0, 5, 10, 15, 20, 25[a]
Population below the poverty line (%)	POVERTY	63%	63, 65, 68, 70, 75
Food safety (items contaminated) (%) (p/a)	FOODSAF	10%	1, 2, 3, 5, 7, 10
Number of pipeline explosions (p/a)	EXPLO	15	2, 5, 7, 9, 11, 15

km kilometers, *p/a* per annum

[a]Indicates percentage change from base level

Based on the preceding research (qualitative and quantitative) the final version of the DCE contained eight attributes. These eight attributes are listed in Table 1.

As is shown in Table 1, we have set base levels for all of the attributes. This means that the DCE will consider policy changes from the current status quo. In order to place the set of attributes in context we provided all survey participants with a detailed description of the attributes that was considered before each interview was completed.

The way we described each attribute used in the DCE is as follows.

Tax: This attribute is defined to be the monthly increase in Nigerian Naira identified as credible and meaningful from the qualitative analysis. Within the final version of the DCE the Tax attribute was framed as follows:

"The Federal Government of Nigeria intends to formulate a policy to regulate the operations of the oil and gas industry. This policy will ensure reduction in environmental hazards and economic hardship caused by the industry in the Niger Delta region. As a result of this policy, the oil and gas companies will be compelled to implement changes that will mitigate environmental hazards, and promote improvement in environmental quality and human livelihood. It is expected that the companies will incur additional costs that will affect their revenue and the amount of tax normally paid to the government by the companies. Hence, the government intends to impose a tax on every individual living within the region including you.

The tax will be collected by a special task force/Committee on environmental protection which will be made up of selected members of the oil producing communities. This task force will ensure proper implementation of government policies and compliance by the oil and gas companies."

In framing the Tax attribute in this way we avoid issues associated with recycling of revenue collected by the Nigerian Federal Government back to the states.

Land occupied by oil and gas pipelines (km) Number of kilometres of the O&G industry pipeline infrastructure within the Niger Delta. The baseline level was established by drawing on data provided by Akpoghomeh and Badejo (2006) and UNDP (2006). The impact of pipeline construction on forest fragmentation and associated biodiversity loss is examined in detail by Agbagwa and Ndukwu (2014).

Unemployment rate in Nigeria (%) The baseline level was established with reference to NBS (2012). Many of those individuals' classified as unemployed might be more appropriately defined as being under-employed (NBS 2015).

Table 2 Example choice card

Attributes	Base level	Alternative options	
	Option A	Option B	Option C
Tax (monthly increase in Naira)	0	N400/month	N100/month
Land occupied by oil and gas pipelines (km)	4500 km	10% reduction	20% reduction
Unemployment rate in Nigeria (%)	20%	30%	35%
Land and water pollution from oil spills	320	60% reduction	50% reduction
Amount of gas flared (billion cubic feet/day)	2.5	15% reduction	10% reduction
Population below the poverty line (%)	63%	68%	70%
Food safety (items contaminated p/a) (%)	10%	2%	1%
Number of pipeline explosions p/a	15	9	11
I would choose [tick one box only]	[ ]	[ ]	[ ]

Number of oil spills per annum The baseline level was set with reference to the number of oil spill cases in the Niger Delta in 2007 (Shell Petroleum Development Company 2012). There is extensive discussion about the nature of environmental damage caused by oil spills resulting from sabotage as well as normal operations (see Anifowose et al. 2012).

*Amount of gas flared (1000 standard cubic feet (SCF) per day)*Anejionu et al. (2015) provides an extensive review of the environmental consequences of flaring from the O&G industry. The extent of SCF per day as a baseline level was estimated to be 2.5 SCF per day drawing on data reported by Omiyi (2001).

Population below the poverty line (%) The baseline level was taken from UNDP (2013). There are regional differences in poverty levels reported in Nigeria with 69% reported for rural areas and 51% for urban areas (NBS 2012). Although levels of poverty have declined in Nigeria the level of economic growth observed since 2000 has not resulted in the reductions in poverty that would be expected.

Food safety (items contaminated (% per annum)) There is scientific evidence linking O&G activities with food safety issues reported in UNDP (2006). Thus, we set the base level at a value above zero informed by the literature and examined reductions from the baseline level.

Number of pipeline explosions (per annum) Pipeline explosions serve as an indicator of poor infrastructure management as well as social unrest (Akpoghomeh and Badejo 2006). The annual number of pipeline explosions vary, as do the number of associated deaths (Anejionu et al. 2015). No specific data recording the number of explosions exists, therefore we took the expected values of focus group respondents to form a baseline value.

3.2.3 Choice Set Design

Given the set of attributes and their associated levels, it was decided to provide three options. The first option in all choice sets was a status quo option. The levels selected for the status quo reflected the information collected during the design stage of the DCE. The two remaining options, both unlabeled, represented a change from the status quo.[6] We then

[6] Each choice card also offered a "don't know option. In total only 30 respondents out of 446 answered "don't know" for one out of eight choice cards, with only two people answering don't know to two out of eight choice cards (34 responses out of 3568 or 0.94% of responses). Given the very low rate of don't know

employed an efficient design assuming a multinomial logit utility specification employing D-error as the measure of selection (Scarpa and Rose 2008). We assumed null priors on our model coefficients. To ensure respondent engagement with the DCE we blocked the 32 choice sets into four groups such that each respondent was asked to complete eight choice sets. The survey questionnaire also contained supplementary questions about the demographics of the respondent as well as debriefing questions. An example of a choice card is presented in Table 2.

3.3 DCE Sample

We aimed to get at least 450 respondents with approximately even numbers in each of the three state. Approximately, 150 respondents were sampled from each of the three states. Within the three states 15 oil producing communities were selected for the study, from five LGA in each of the states:

- *Akwa Ibom State* Edo (Esit Eket LGA), Iko (Eastern Obolo LGA), Mkpanak (Ibeno LGA), Unyenge (Mbo LGA), Ukpene Ekang (Ibeno LGA).
- *Bayelsa State* Odi (Kolokuma-Opokuma LGA), Imiringi (Ogbia LGA), Etiama (Nembe LGA), Okotiama-Gbarain (Yenagoa LGA), Ogboibiri (Southern Ijaw LGA)
- *Rivers State* Chokota community (Etche LGA), Igbo-Etche (Etche LGA), Alesa-Eleme (Eleme LGA), Obigbo (Oyigbo LGA), Biara (Gokana LGA).

This sampling strategy meant that 30 respondents were selected from each of the 15 communities. We recruited respondents door-to-door using a systematic sampling approach. Thus, every third house was contacted after a random starting point between the first and tenth house in each community. Not all houses visited agreed to participate in the survey, and not all respondents were suitable to participate i.e. all respondents need to be at least 18 years of age and we also required at least a third of respondents to be female. This was because the questionnaire was designed to collect individual responses, not the opinions of a household or group. In total 455 respondents were successfully contacted and agreed to participate. Once a household agreed to participate, data collection was undertaken face-to-face with a member of the household by trained field assistants. A total of 446 respondents completed the questionnaire and were used in the subsequent analysis. The descriptive statistics of the respondents are provided in Table 3.

In Table 3, it can be seen that the samples are drawn almost equally from the three states. However, male participation was more easily obtained. In order to reduce the number of male respondents, we required, as noted that on average at least every third respondent was female. The impact of this sampling strategy can be see by the percentage of female respondents in our sample. However, despite this approach to data collection, our sample has more male respondents than would be expected given the gender mix in Nigeria.

If we consider family size, we see that there are some small differences between the states. Drawing on available Nigerian Federal Government statistics (NBS 2012, 2015) it

Footnote 6 (continued)
responses and the fact that the survey included a status quo option we did not include these responses in the analysis presented.

Table 3 Survey respondents descriptive statistics

Region	Akwa Ibom	Bayelsa	Rivers	Full
Sample size	149	148	149	446
Gender				
Male	100 (67)	94 (64)	96 (64)	290 (65)
Female	49 (33)	54 (36)	53 (36)	156 (35)
Martial status				
Single	53 (35)	35 (24)	46 (31)	134 (30)
Married	86 (58)	104 (70)	92 (62)	282 (63)
Widowed	10 (7)	9 (6)	11 (7)	30 (7)
Education				
Non-formal	23 (15)	19 (13)	22 (15)	64 (14)
Primary	27 (18)	13 (9)	17 (11)	57 (13)
Secondary	49 (33)	49 (33)	52 (35)	150 (34)
Tertiary	50 (34)	67 (45)	58 (39)	175 (39)
Family size				
1–3	63 (42)	57 (39)	66 (44)	186 (42)
4–6	63 (42)	68 (46)	58 (39)	189 (42)
7–9	22 (15)	21 (14)	24 (16)	67 (15)
> 10	1 (1)	2 (1)	1 (1)	4 (1)
Occupation				
Farming	27 (18)	18 (12)	21 (14)	66 (15)
Civil servant	18 (12)	36(24)	21(14)	75(17)
O&G	11 (7)	11 (7)	10 (7)	32 (7)
Other	11 (7)	4 (3)	12 (8)	27 (6)
Self employed	33 (22)	41 (28)	42 (28)	116 (26)
Unemployed	10 (7)	19 (13)	21 (14)	50 (11)
Student	31 (21)	14 (9)	21 (14)	66 (15)
Fishing	8 (5)	5 (3)	1 (1)	14 (3)
Income (N)				
< 10,000	56 (38)	41 (28)	61 (41)	158 (35)
10,000–20,000	37 (25)	46 (31)	39 (26)	122 (27)
21,000–30,000	20 (13)	23 (16)	19 (13)	62 (14)
31,000–40,000	12 (8)	7 (5)	5 (3)	24 (5)
41,000–50,000	8 (5)	4 (3)	6 (4)	18 (4)
51,000–60,000	5 (3)	5 (3)	4 (3)	14 (3)
61,000–70,000	4 (3)	4 (3)	3 (2)	11 (2)
>71,000	7 (5)	18 (12)	12 (9)	37 (8)
Age (years)				
Below 20	6 (4)	1 (1)	39 (2)	10 (2)
20–29	28 (19)	18 (12)	25 (17)	71 (16)
30–39	35 (23)	54 (36)	37 (25)	126 (28)
40–49	48 (32)	47 (32)	48 (32)	143 (32)
50–59	29 (19)	24 (16)	31 (21)	84 (19)
60 and above	3 (2)	4 (3)	5 (3)	12 (3)

is reported that average household size in 2009/2010 was 4.2 in Akwa Ibom, 3.7 in Bayelsa and 4.5 in Rivers. If we compare these to our sample then we see that we have sampled slightly larger households than average in Bayelsa whereas for the other two states we have replicated the state level average closely.

Martial status participation also shows some variation between the states. Comparing these to official data we find for Akwa Ibom that 47.3% are married and 41.4% never married. For Bayelsa we have 70% married compared to 54.4% and those single 24% compared to 33.8%. Finally, for Rivers we have 62% married in the sample compared to 49.3% and for single we have 31% compared to 43.1% from official data. Thus, in terms of marital status we have a slightly higher representation of married individuals than is typical.

In terms of the level of educational attainment of our sample, our respondents are significantly more educated than Nigerians on average. For example, only 14.9% of Nigerian's attained post-secondary education 2008 compared to our sample average of 39%.

Finally, comparing our income data to state level information in 2009/2010 for Akwa Ibom the NBS (2012) survey reports that 78.8% households have income less than 20,000 Naria per month. For our sample we have 63% on average with Bayelsa at 59% whereas official data reports 61.5%. For Rivers our sample has 67% with official estimates of 56.1%. In summary, although our sample do not exactly correspond to the official statistics, our sample is broadly representative.

4 Model Priors and the Number of Mixtures

4.1 Model Priors

Before we present our results we explain our choice of priors and the impact on model estimates. In particular, we examine how our choice of specific priors influenced the number of mixtures produced by the BIML. Importantly, for our DCE data, we found some substantive differences, although in all important respects, we could have presented the results for any of the models considered and drawn the same basic conclusions regarding our DCE.

We arrived at our initial set of priors as follows.

First, our selection of G_0 was 'Empirically Bayesian', in that we estimated a standard fixed parameter Logit (FPL) (under non-informative priors) in the first instance. The largest (in absolute terms) of the parameters within the fixed parameter model was a value of just below six. Since the base distribution should easily encompass the parameters in a fixed parameter model, we specified $G_0 = N\left(0, g_0 \times I\right)$ where we experimented with three values for the standard deviations for base distribution: $\sqrt{g_0} = (6, 10 \text{ and } 15)$. Alterations in the base distribution did have a larger effect on the results. This was most evident with the least diffuse base distribution standard deviation $\sqrt{g_0} = 6$, which tended to push the marginal utilities towards zero. The effect was less obvious between the more diffuse distributions $\sqrt{g_0} = 10, 15$. Importantly, however, while each of the individuals estimates were changed in terms of there overall magnitudes, the rankings of marginal utilities for individuals remained very similar for each of the attributes, and the WTP estimates even more similar. In what follows, we present (mainly) the results for the middle distribution $\sqrt{g_0} = 10$.

Second, for estimation, we experimented with three sets of hyper parameters for the mean and variance for the Gamma distribution for α. The priors on the mean and variance of the concentration parameters had virtually no effect (at both the individual and

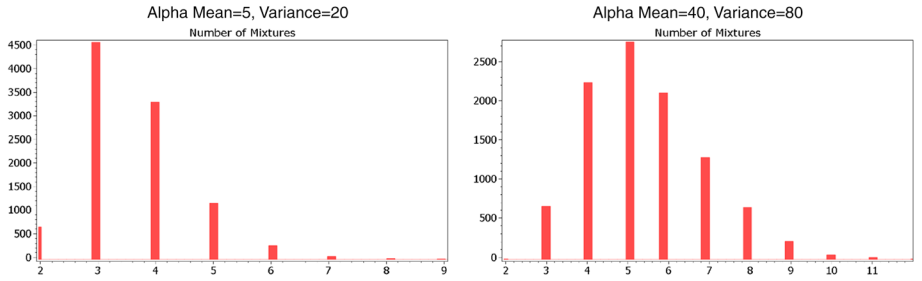

Fig. 1 Frequency of number of classes under different priors

average level) at $E(\alpha) = 5$ with $Var(\alpha) = 5$ or $Var(\alpha) = 20$, when comparing between any two models with the same base distribution G_0. Thus, the more diffuse distribution did not have any tendency to increase the number of mixtures that were required to model the data. However, when using the more informative prior ($E(\alpha) = 40, Var(\alpha) = 80$) this increased the number of mixtures and there was a moderate change in the individuals' distributions, such that they became more "smooth", with a greater number of mixtures being introduced. Importantly, this did not change the results dramatically with regard to the marginal utility estimates for individuals. Thus, the results we present are essentially the same regardless of the settings for the hyper parameters for the concentration parameter. Therefore, the model results we present are for the most diffuse prior distribution for α with $E(\alpha) = 5$ and $Var(\alpha) = 20$.

4.2 Model Classes

We now consider how the number of classes (or mixtures) are impacted by the choice of priors. As explained, new classes are introduced and eliminated as part of the estimation process. The number of classes at each of the 10,000 sample points were also recorded. We undertook this exercise for our preferred set of priors and the more diffuse set of priors for the concentration parameter (α). The frequency of the number of classes in both cases is shown in Fig. 1.

In Figs. 1 and 4a is for the relatively less diffuse prior with mean five, were as Fig. 4b is for the more diffuse prior and as such favours less concentration. As can be seen in Fig. 4a, the mode is three classes with four and five classes being more common than two. At no point during the estimation procedure were all individuals placed in the same mixture (which is indicated by the zero frequency of one mixture only), though there is no restriction against this happening. However, it should not be surmised on the basis of these results that there are three or four distinct groups, since if there is a continuous distribution of individual preferences, this will induce multiple mixtures.

Turning to Fig. 4b, when we employ a more diffuse prior we observe an increase in the estimated number of classes. For Fig. 4b the mode is five, with seven or eight classes also being relatively common, and two extremely rare. Readers might be tempted to assume that this will have a substantial impact on the resulting estimates especially with to WTP. However, the dependence of this prior turns out far less important than one might expect given how the resulting parameter estimates and associated WTP are arrived at.

Table 4 Means and standard deviations for marginal utilities

Attribute	BIML				HBL				FPL	
	$\hat{\mu}$	$Sd(\hat{\mu})$	$\hat{\sigma}$	$Sd(\hat{\sigma})$	$\hat{\mu}$	$Sd(\hat{\mu})$	$\hat{\sigma}$	$Sd(\hat{\sigma})$	$\hat{\mu}$	$Sd(\hat{\mu})$
Tax	0.63	0.10	0.86	0.33	0.50	0.03	0.39	0.15	0.39	0.02
Landocc	1.07	0.46	0.70	0.71	1.28	0.30	1.50	1.01	0.79	0.25
Unemp	3.36	0.83	1.16	1.13	4.26	0.72	1.49	1.44	3.73	0.62
Spill	1.11	0.40	0.29	0.61	1.48	0.24	1.35	0.95	1.12	0.19
Flare	0.46	0.56	0.79	0.97	0.11	0.36	0.74	0.69	0.19	0.31
Poverty	4.83	0.92	0.33	0.95	6.93	0.96	1.17	1.37	5.01	0.73
Foodsaf	5.06	1.39	1.03	1.22	6.55	1.53	0.96	1.199	5.68	1.28
Explo	1.38	0.34	0.17	0.57	1.87	0.19	1.67	0.85	1.41	0.13
Statquo	− 3.14	0.45	1.32	0.92	− 3.62	0.27	0.57	0.57	− 2.56	0.19

$\hat{\mu}$ = Mean Marginal Utility, $Sd(\mu)$ =Standard Deviation of μ

$\hat{\sigma}$ =Standard Deviation of Marginal Utilities (BIML and HBL only)

Finally, neither Fig. 4a, b reveal how different the results within each component are. We can, however, reveal this information by examining the resulting marginal utility distributions which can be seen in Fig. 2 which we present in the next section.

5 Empirical Results

In the results presented, we compare the BIML results with the HBL (all parameters being specified as normally distributed) and a standard FPL. All models have been estimated employing Bayesian methods. The estimation of the HBL is now reasonably standard with specific details covered in Train (2003) so we do not repeat this here.

5.1 Model Estimation and Convergence

Our BIML model specification was estimated using the DP algorithm described in Sect. 2. To generate the results, we report, we undertook a 'burn-in' of 50,000 draws. In this case a draw means a complete update of all (re)assignments and parameters having cycled completely through the individual estimates for each of the respondents. We then undertook a further 500,000 draws from which each 50th draw was collected (this is known as thinning the sample). The took this approach to reduce the likelihood of serial correlation being an issue. This yielded 10,000 draws from which to estimate the mean and variance parameters for the posterior. The initial number of classes (m) was set at two, but the sampler rapidly changed from this value within a few iterations and the results are independent of this initialisation. For the HBL models, we also took every 500th draw so as to obtain an accurate and relatively uncorrelated mapping of the posterior distribution.

To assess overall model convergence we considered visual plots of the sequence of data draw as well as undertaking formal tests for stability of the MCMC estimates. Gelman and Rubin (1992) Rhat was applied to two sets of runs initialised randomly at different starting points. The Gelman-Rubin Rhat delivered values of less than 1.01 for all parameters, which is much less than the recommended maximum value.

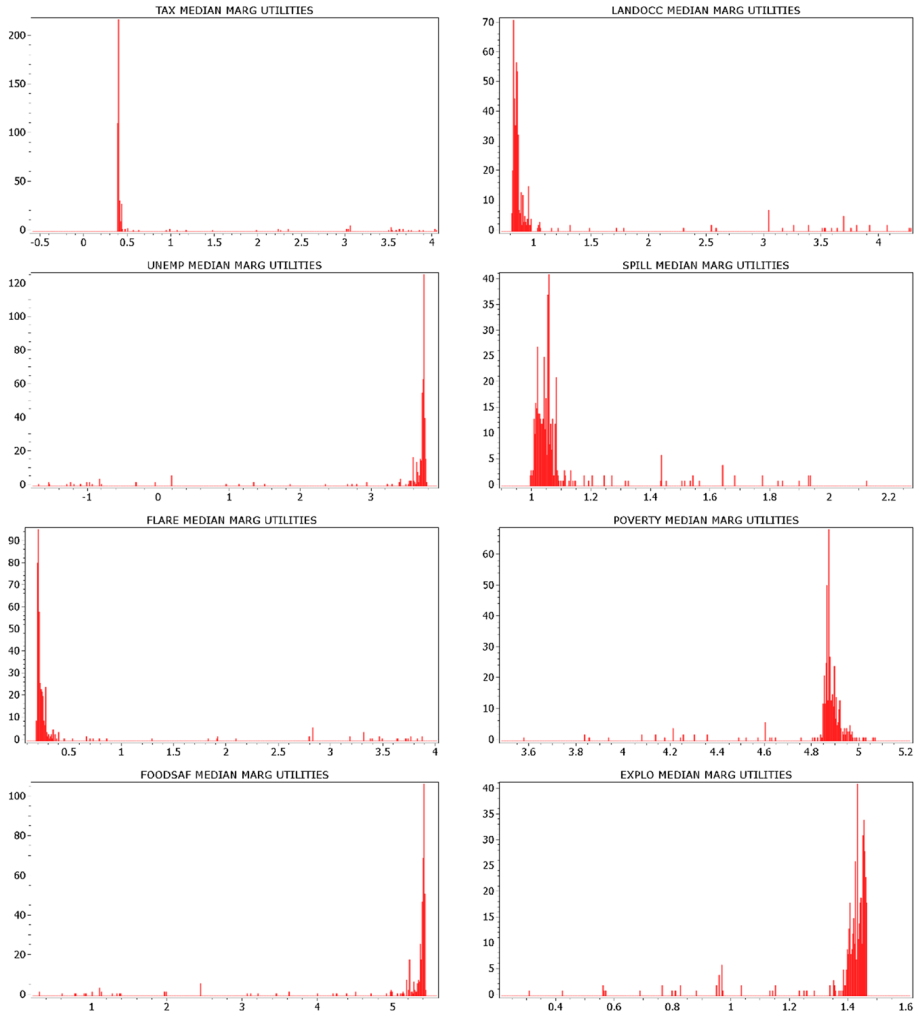

Fig. 2 BIML estimates of individuals marginal utility by attribute

5.2 Marginal Utility Estimates

Prior to estimation, we made sure that the signs and scales of the attribute variables were such that we could more easily check whether the parameter estimates were consistent with our a priori expectations. TAX was, therefore, entered as a negative value so that its coefficient would be expected to be positive. The model was parameterised so that with regard to the UNEMP, POVERTY, and FOODSAF parameters, the resulting coefficients are marginal utilities for a percentage change for these quantities (e.g., the value of a 1% shift in unemployment from 21 to 20%). In contrast, the coefficients for LANDOCC, SPILL, FLARE, and EXPLO represent the marginal utilities for a percentage change relative to the base level (e.g., the value of a 1% fall in Flares from 2.5 billion cubic feet per day).

Table 5 WTP estimates from the BIML

Attributes	1% reduction from	Mean	95%	Median	5%
Landocc	4500 km	2.11	2.30	2.20	1.10
Unemp	20%	8.47	10.13	9.71	0.03
Spill	320 pa	2.49	2.85	2.73	0.69
Flare	2.5 SCF/day	0.70	1.32	0.57	0.45
Poverty	63%	11.36	13.22	12.07	1.65
Foodsaf	10% items contaminated pa	12.44	14.7	14.18	0.72
Explo	15 pa	3.29	3.96	3.74	0.40
Statquo	–	6.56	7.48	7.18	2.45

Table 4 presents the mean[7] and standard deviation of the individual marginal utilities for the BIML in the first two columns. For reference, the estimates for the FPL and HBL are given in the columns to the right.

From Table 4, we can see that the estimates of the BIML means are similar to the FPL and HBL models. In and of itself, Table 4 is not particularly informative since it gives only a partial indication as to the distribution of the marginal utilities. In order to get a better sense of the distribution of marginal utilities, we turn to Fig. 2 which has histograms for the BIML estimates of individuals marginal utilities.

In the top left hand corner of Fig. 2, we have the marginal utility for the (-)TAX option. As can be seen, the mean for this attribute (0.65) reflects a very partial picture. While we have a spike just below 0.5, for individuals with relatively small negative utility for TAX increases, there is another dispersed distribution for individuals who have a much larger negative utility for TAX. This does not necessarily indicate very much about preferences, since the differences in marginal utilities may reflect scale effects rather than differential propensities to trade-off attributes. However, we can contrast the marginal utilities for TAX with the marginal utilities for variables such as EXPLO and FOODSAF which have a greater mass of individuals at the upper end of the ranges.

Our general observation is that these distributions are certainly not bell shaped and therefore not of the type that are easily captured by normals such as in the case of the HBL. For the other two base distributions that we used, the same patterns were observed, except that the distributions were more compact for the tighter base distribution (G_o) and more spread out for the more diffuse base distribution. Thus, commonly employed distributions within a HBL specification, such as the log-normal would fail to capture the distributions for EXPLO or FOODSAF. Equally, however, there is little sign of bimodality or multimodality in the marginal utility distributions that suggest clear groupings, with the vast majority of individuals clustered close together (around 90% of the sample) with another 10% being widely distributed away from this majority point.

5.3 WTP Estimates

The marginal WTP for an attribute is calculated as the marginal utility for the attribute divided by the marginal utility of the TAX variable. In running the BIML model, no

[7] Note that a mean of medians or median of median estimates are almost the same.

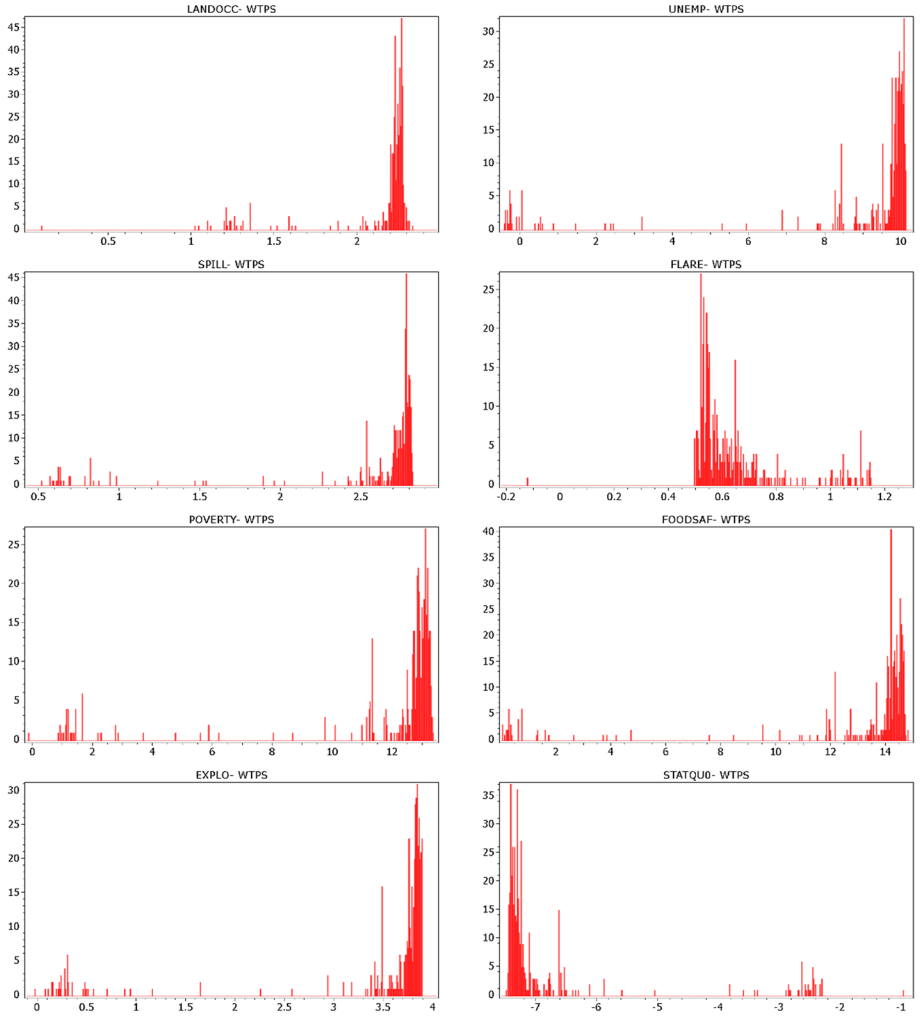

Fig. 3 Individual willingness to pay estimates from BIML

explicit inequality constraints were placed on the parameters. Thus, the marginal utility distribution for the TAX attribute could have mass close to zero or be negative. Without further constraints the first and second moments for the WTPs will not necessarily be finite, and the empirical means and variances may be dominated by a few observations where the denominator is close to zero, resulting in very high WTPs. Accordingly, we censored the marginal utility distributions for each individual so that marginal utilities were accepted only if they generated WTPs less than N100 per month for a 1% change in the attributes.[8] These represent weak constraints in the sense that the extreme values exceed any realistic

[8] In principle, we could impose a prior on our model to achieve the same result. However, to do so would require we first estimate the model as there is little reason a priori to assume such outcomes. As such it makes practical sense to impose this restriction ex post.

WTP and the distributional mass in these regions was low. Implementing constraints twice as stringent or twice as lenient resulted in almost identical WTP distributions.

The results presented in Table 4 are individual's WTPs for each of the attributes. Specifically, we give the mean, median, lower and upper percentiles for individual's WTPs.

As we can discern from Table 5, nearly all individuals have negative WTPs for the status quo and positive WTPs for reductions in the "bads" associated with O&G production. These estimates represent WTPs for a 1% reduction for each attribute with the exception of the status quo. For example, the mean WTP for a 1% reduction in LANDOCC (from 4.500km) is N2.15 per month, and a 1% decrease in contamination (FOODSAF) to 9% (from 10%) is worth on average N12.77 per month. Overall, these results indicate that respondents generally believe that some sort of action should be taken for the mitigation of negative effects of O&G activity. However, the lower 5% column reported in Table 5 shows that the average and median figures obscure the fact that there are a small but substantive mass of individuals at the lower end of the distribution who are prepared to pay very little. This feature of the results is also demonstrated in Fig. 3.

From Fig. 3, we can see that for seven out of the eight attributes we have a greater mass packed around an upper value (or lower value in the case of the STATQUO) with a more sparsely packed distribution towards zero. For five attributes there appears to be a small amount of bunching around or close to zero. The FLARE attribute is a notable exception with the greater mass packed at the lower end of the WTP range.

It is also worth noting the differences in interpretation of UNEMP, POVERTY, and FOODSAF coefficients relative to the others. These three are WTPs for changes in total percentages rather than as percentages relative to the baseline. Thus, a 1% change in unemployment is equivalent to a 5% increase relative to its baseline level of 20%. Therefore, while it is evident that the WTPs in relation to the UNEMP, POVERTY, and FOODSAF attributes are large relative to the others, one should not conclude that the WTPs for these attributes is unambiguously greater. In addition, it is clear that individuals are willing to pay significant amounts to reduce LANDOCC, SPILL and EXPLOSIONS.

Of maybe greater importance is that these results indicate that respondents are not prepared on average to sacrifice 1% higher unemployment or poverty in order for a 1% reduction in LANDOCC, SPILL and EXPLOSIONS, relative to their baseline levels. This result indicates that if the O&G industry is able to contribute to reductions in poverty and unemployment, then our respondents are willing to bear some of the negative environmental effects of the O&G industry.

5.4 Distinct Groups

The nature of the plots in Figs. 2 and 3 do not clearly indicate multiple groups. Rather, they illustrate that there is a mass of individuals clustered around a common point, with a minority distributed away from this to varying degrees, yet not obviously forming an alternative cluster. To some, the rather "untidy" plots will hinder a simpler but more popular way (albeit imprecise) to report DCE results, which is to report only means and variances. However, the plots also reveal that employing a FM approach to analyse the data will fail to reveal to the true nature of the "clusters" that may well be important when it comes to developing a more thorough understanding of preferences.

One way to sharpen our interpretation is to conduct a cluster analysis on the WTP estimates for all eight attributes. A drawback of the Bayesian approach relative to the classical approach to mixtures is that the Bayesian approach has no natural

Table 6 WTP estimates by cluster

Attributes	1% reduction from	Mean	Cluster 1	Cluster 2
LANDOCC	4500 km	2.11	1.30	2.21
UNEMP	20%	8.47	0.72	9.43
SPILL	320	2.49	0.88	2.69
FLARE	2.5 billion cubic feet/day	0.70	1.01	0.66
POVERTY	63%	11.36	2.70	12.43
FOODSAF	10% items contaminated p/a	12.44	1.74	13.76
EXPLO	15 pa	3.29	0.63	3.62
STATQUO	–	6.56	2.74	7.04

Cluster 1 n = 49, Cluster 2 n = 397

classification of groups in the sense that each individual is assigned to a particular class because the classes themselves can "label switch" and evolve. This is not only an issue associated with the infinite mixtures approach but also pertains to the Bayesian finite mixtures approach. Our purpose here is not to categorise respondents. However, to highlight the bimodal nature of the utilities and estimates we perform a k-means cluster approach. If a k-means cluster analysis is then conducted either on WTPs estimates, around 10% of the sample is placed in one group and 90% in the other and the cluster means of these two groups tend to be very different. If more than two groups are specified then the cluster means of two of the groups tends to be very similar with another clearly differentiated from the others.

We present the clustered means for the WTPs for the k-means of the two groups in Table 6.

What we observe from Table 6 is that the large cluster has (unsurprisingly) WTP estimates that are very close to the overall means. However, those differentiated from this cluster have much lower WTPs for all the attributes except for the FLARE attribute which is still relatively small. Therefore, we have seem to have one large relatively homogenous group, with substantial WTPs that contain around 90% of individuals. There is another distinct 10% who are behaviorally very different. They appear to have very low WTPs for all the attributes.

5.5 In Sample Performance

As stated above, we do not attempt to perform a proper model comparison for the data set above. Such an excercise would require a proper comparison of appropriately penalised model criteria requiring either the Bayesian information criteria or Marginalised Likelihoods. We did, however, examine the simple predictive performance of the BIML and HBL using the average estimates for each of the respondents. The BIML model predicted approximately 73% of choices correctly, with the HBL doing slightly better (at approximately 75%). Naturally, this is not a result that will convince readers that the BIML is the superior approach. However, we do not believe the viabilty or lack thereof of a given modelling approach should not be judged based on a singular data set. Additionally, we would note that using the BIML approach there were a number of individuals that whose responses were far better characterised using the BIML approach, and others worse. As for the parameters themselves, the BIML tended to have a greater bunching around given

points of predictive accuracy than the HBL approach. Thus we would argue that using this crude measure the BIML did about as well as the HBL overall, while giving quite different estimates for marginal utilities.

6 Discussion and Conclusions

In this paper, we have presented the results of a DCE undertaken in the Niger Delta region of Nigeria in order to better understand peoples WTP for the mitigation of negative impacts associated with the O&G industry. The DCE examines respondents' valuations of environmental goods and food security but also considers changes in poverty and employment. It is the first study to use a DCE within this context. To model the data we employed a 'Bayesian Infinite Mixtures Logit' (BIML) approach in order to model heterogeneity in preferences. This approach allows for a potentially large number of mixtures, but does not require the researcher to fix or decide on the number of mixtures. This approach offers a flexible alternative to either a finite mixture approach or the popular 'Mixed Logit' or 'Hierachical Bayes' approaches that require well specified continuous distributions.

When applied to our DCE data, the BIML had slightly worse in-sample fit relative to the HBL but revealed that there was a tendency towards bimodal WTP distributions with most (around 90%) respondents within our sample willing to pay considerable amounts for a number of attributes and a minority (around 10%) being prepared to pay very little. The distributions which emerged under the BIML approach were quite different to those that would be expected under a standard HBL model employing standard distributions (e.g. normal or log-normal distributions). However, our data here could have been (retrospectively) modelled well by a finite mixture model with as few as 3, but possibly as many as 5 classes.

The BIML nests alternative model specifications (FPL, HBL and FM). We emphasise that the BIML can do a good job at capturing both continuous distributions and/or a small number of classes (mixtures). Currently, researchers tend to employ one of the approaches listed above, without any real justification for their specific choice (ex ante or ex post). Indeed, the BIML model can be viewed as an alternative to the classical LML introduced by Train (2016) and examined in a series of recent papers. Like LML, the BIML does provide the researcher with far greater flexibility with regard to how we model heterogeneity of preferences. There is no doubt that these new approaches to modelling DCE data warrant further attention. However, it should not be assumed that the HBL or FM is no longer an appropriate model choice. But, as we have demonstrated with our case study (and appendix), the BIML warrants increased use and attention in future research because of its ability to capture a wide array of data generating processes that can be incorrectly modelled when making standard parametric assumptions in model implemetation.

Turning to the specific policy issues examined by the DCE, while the O&G industry provides both employment and wealth, these benefits are not always distributed widely or equitably. Interestingly, the Nigerian Federal government allocates state funding on the basis of population size not revenue generation (Imobighe 2011). Thus, the wealth created by the O&G sector, has not always benefited O&G regions. Accordingly, the Nigerian Federal government has passed legislation allowing oil rich states to obtain 13% of the revenues from mineral extraction. But the percentage amount has varied significantly over the period of O&G activity, because of political tensions between the northern and southern states in Nigeria (Songi 2015).

Importantly, the public frequently view the O&G industry as a competitor for key resources and as a result a contributor to the poor economic and environmental situation of many people, rather than a source of wealth and prosperity for all. However, despite this context, our DCE results indicate that respondents in the Niger Delta region are willing to pay significant amounts to mitigate the 'bads' associated with the O&G industry. This is because, our respondents also value the benefits that come with higher levels of employment and poverty reduction. When taken together the results indicate that preferences for maintaining employment and minimising poverty are at least as important as reducing environmental externalities. Furthermore, the DCE revealed that improvements in food safety is the most highly valued attribute which places the importance on meeting the basic human needs of society in context. These results are not surprising when we consider that some 60% of the Nigerian population are considered to be living in poverty (UNDP 2015).

Our findings might be partly explained with reference to an environmental Kuznets curve whereby, in the states we studied, the turning point where economic activity at which environmental protection begins to gain greater value compared to economic development has yet to be achieved. Therefore, at least as far as the problems considered in this research are concerned, policies that attempt to deal with the environmental consequences of O&G activity will need to be implemented in such a way that they do not impact O&G contributions to employment and poverty reduction. That said, our results do reveal that people are generally prepared to pay for the mitigation of negative environmental effects. This does not imply that the public should pay for mitigation especially if Nigeria is to frame environmental policy in terms of the polluter pays principle. However, we did not observe any tendency towards widespread refusal to pay for mitigation nor a sense of indignation at the prospect. The WTP estimates provided here may give some insights as to which impacts should be minimised. However, any policy would need to take account of the relative costs of alternative forms of mitigation not considered here.

Finally, the DCE policy scenarios assume that collected tax revenue will go directly to a local task force that will use the funds accordingly. As noted, this would be a change in how government funding works in Nigeria. Respondents appeared to be comfortable with this funding model. The need for alternative ways to manage and share the economic benefits that arises as a result of O&G activity suggests the need for greater LGA and local community involvement. This research provides evidence that respondents are comfortable with the potential use of mechanisms such as the foundation, trusts and funds model advocated by Songi (2015). The potential for this type of benefit sharing model is a topic for further research.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix 1: Demonstrating the Flexibility of the Infinite Mixture Approach

Introduction

The purpose of this appendix is to demonstrate the flexibility of the BIML. Specifically, we demonstrate that the BIML approach has the desirable attribute of being able to deal with small and large dimensional mixtures, even when the underlying model has coefficients that

come from a continuous distribution. To this end we present a simple monte carlo exercise examining relative model performance.

We use two polar data generating procedures, one which is continuous and wholly consistent with a HBL model (in which case the BIML is "mispecified") and another where there a finite mixture (in which case the HBL is "mispecified"). The expectation is that the HBL and BIML will outperform the other when the data generating process (DGP) is specified in a way that is consistent with these models. However, of more interest is how they perform when used to model a data from a DGP that is not wholly consistent with each of the respective models.

A key step in understanding the BIML approach is to appreciate that it is not the central goal of this approach to select the mixture dimension. The BIML will be outperformed by a finite mixture model using a DGP which is finite, providing the number of mixtures is known. Practitioners that use latent class models commonly view the selection of the number of classes as a specification issue rather than seeing the number of classes as a parameter to be estimated (along with errors classifying the uncertainty around these estimates). By contrast, the BIML views the number of classes as a parameter of the model, albeit a discrete one. Within the Bayesian setting it has a distribution and the uncertainty about the number of classes is quantified within the estimation procedure, and reflected in the distributions of other parameters in the model.

Below, we generate two data sets from two distinctly different DGPs. One is a linear discrete mixture with three sets of fixed utility coefficients each representing a class, and the other is a continuous normally distributed set of utility coefficients, in accordance with the assumptions underpinning a standard HBL. These two data sets are then each estimated using the HBL and the BIML and the results compared. For the discrete mixture data, we also compare the BIML with the results obtained by specifying exactly three classes. We outline this in more detail below.

The Data Generating Processes

The data used herein is generated from the process

$$U_{ijs} = x'_{ijs}\beta_j + e_{ijs}$$

where j is the respondent, and x'_{ijs} represents the vector of attributes under the ith option in the sth choice set given to the jth respondent (here a draw from a standard normal vector of independent components). The vector x'_{ijs} is a one by five vector, thus representing five attributes within a discrete choice experiment (DCE). Each error e_{ijs} is Gumbel distributed independently across i, j and s. The parameters β_j are "marginal utilities" (MUs). As is standard in discrete choice models, it is assumed that individuals choose the option (i) with the highest utility.

For the demonstration below each respondent makes 10 choices ($s = 1, \ldots, 10$), where each choice set has three options ($i = 1, 2, 3$), and, there are 100 respondents (thus 1000 choices in total). The two DGPs are identical and are only distinguished by the way the vectors β_j are generated.

- *Case 1* For the continuously distributed preferences we specify:

$$\beta_j \sim N\left(b^1, I\right) \tag{7}$$

$$b^1 = (1, 1, 1, 0, 0) \tag{8}$$

- *Case 2* For the discrete mixtures case β_j is randomly selected (with equal probability) from one of b^1, b^2, b^3 below:

$$\begin{aligned} b^1 &= (1, 1, 1, 0, 0) \\ b^2 &= (-1, -1, -1, 0, 0) \\ b^3 &= (0, 0, 0, 1, 1) \end{aligned} \tag{9}$$

Below we will refer to these two distinct cases as the continuous DGPs (Case 1) and discrete (Case 2). For Case 1 we are examining how well the BIML performs when the underlying DGP is accurately captured by the HBL (and is inconsistent with any finite mixture). For Case 2 we are examining how well the BIML performs without knowledge of the correct number of mixtures, and how well the HBL performs when the assumption of a continuous distribution is clearly false. Importantly, readers should realise that while the HBL assumes an a priori continuous distribution (e.g. normal), the latent estimates for individuals need not conform closely with this distribution.

Estimation Results

Each model is estimated using the HBL and the BIML, using priors that are commonly employed within the literature for the HBL (a Standard Normal for the mean and Wishart priors for the covariance matrix), and priors as in those within the main body of the paper for the BIML.

Distribution for "Mixture Dimension"

We first examine the distribution for the mixtures dimension for the two cases for the BIML as shown in Fig. 4. For the graph on the left hand side ,the true number of mixtures was three, as specified for the discrete DGP (Case 2 above).

Figure 4 shows the distribution for the number of mixtures (NOMs) which as discussed above is a discrete parameter within the BIML model. As can be seen by Fig. 4a (on the left hand side), the BIML selects three mixtures (the correct number) approximately 50% of the time, but also visits four or five mixtures occasionally, and a higher number of

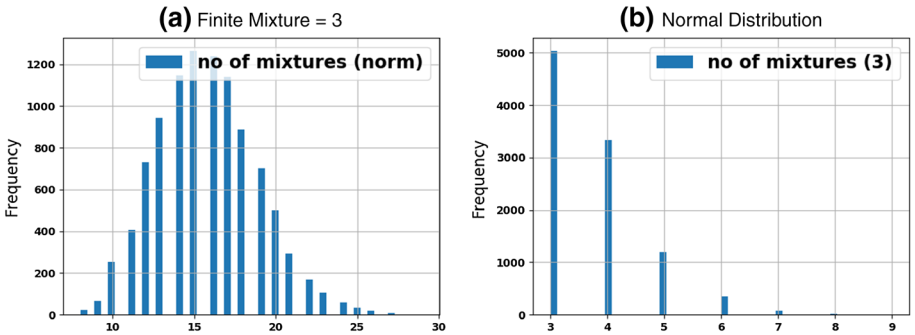

Fig. 4 Distribution of mixture dimension for BIML

Table 7 R^2 for MUs estimates continuous case

	MU1	MU2	MU3	MU4	MU5
HBL v True	0.56	0.58	0.67	0.58	0.65
BIML v True	0.51	0.51	0.58	0.56	0.55
HBL v BIML	0.94	0.91	0.91	0.89	0.89

mixtures more rarely. In this case it never visits less than 3 mixtures. It should be noted that one or more of these mixtures may contain very few respondents, often containing only single individuals, and only exist for short periods.

By contrast, for the continuous DGP (Case 1 above), Fig. 4b (on the right hand side) shows that the number of mixtures generated is very large, the median being 16 or 17. This illustrates that the BIML "mimics" the continuous case by selecting a very large number of mixtures.

Estimates of Individual Marginal Utilities

Case 1: Normally Distributed MUs (i.e. HBL is the "Correct" Model, BIML is the "Incorrect" Model)

We now examine model performance by looking at the individual estimates for respondents. In Fig. 5, we start by plotting each of the estimates of the utility coefficients against there true (Monte Carlo generated) values under the first DGP (Case 1 above). In these circumstances the HBL (or mixed logit) is bound to give the optimal estimates since it is the "correct" model. In Fig. 5. the true values are referred to as beta1, beta2, etc. and are on the horizontal axis. The HBL are labelled as betaml1, betaml2 etc., whereas the BIML estimates are labelled betamx1, betamx2 etc.

As demonstrated by Fig 5a, while there is a good correlation between the actual marginal utilities (from a continuous normal distribution) and their estimated values for all five attributes using the HBL, these estimates still retain considerable noise (unsurprising given that each respondent makes 10 choices). For Fig. 5b when using the BIML we see that, from a visual perspective, the estimates look to have a broadly similar relationship with their true values as compared with the HBL in Fig. 5a. Indeed, while the BIML is bound to be slightly less accurate than the HBL, this is not visually evident for the case in hand. However, as reported in Table 7, in each case the R^2 between the HBL estimates and the true marginal utilities is consistently higher than for the corresponding R^2 for the BIML and the true values. Thus, the results on one hand confirm the obvious. The HBL will outperform the BIML for a correctly specified DGP. On the other hand the BIML seems to do surprisingly well for a discrete mixture model.

But, reassuringly, from Fig. 5c, the relationships between the BIML and HBL estimates is closer than for each of these with the true estimates (in all cases an R^2 for HBL and BIML estimates is around 0.9 in all cases). This reflects the fact that the estimates for the BIML and HBL reveal the signal from the choices made by respondents, not just the underlying values underpinning these choices. Thus, these results demonstrate that when data is generated in accordance with the assumptions behind the HBL, the HBL unsurprisingly does best. However, the performance of the BIML is not far behind.*Case 2: Discretely Distributed MUs (i.e. The BIML is the "Correct" Model, HBL is the "Incorrect" Model)*

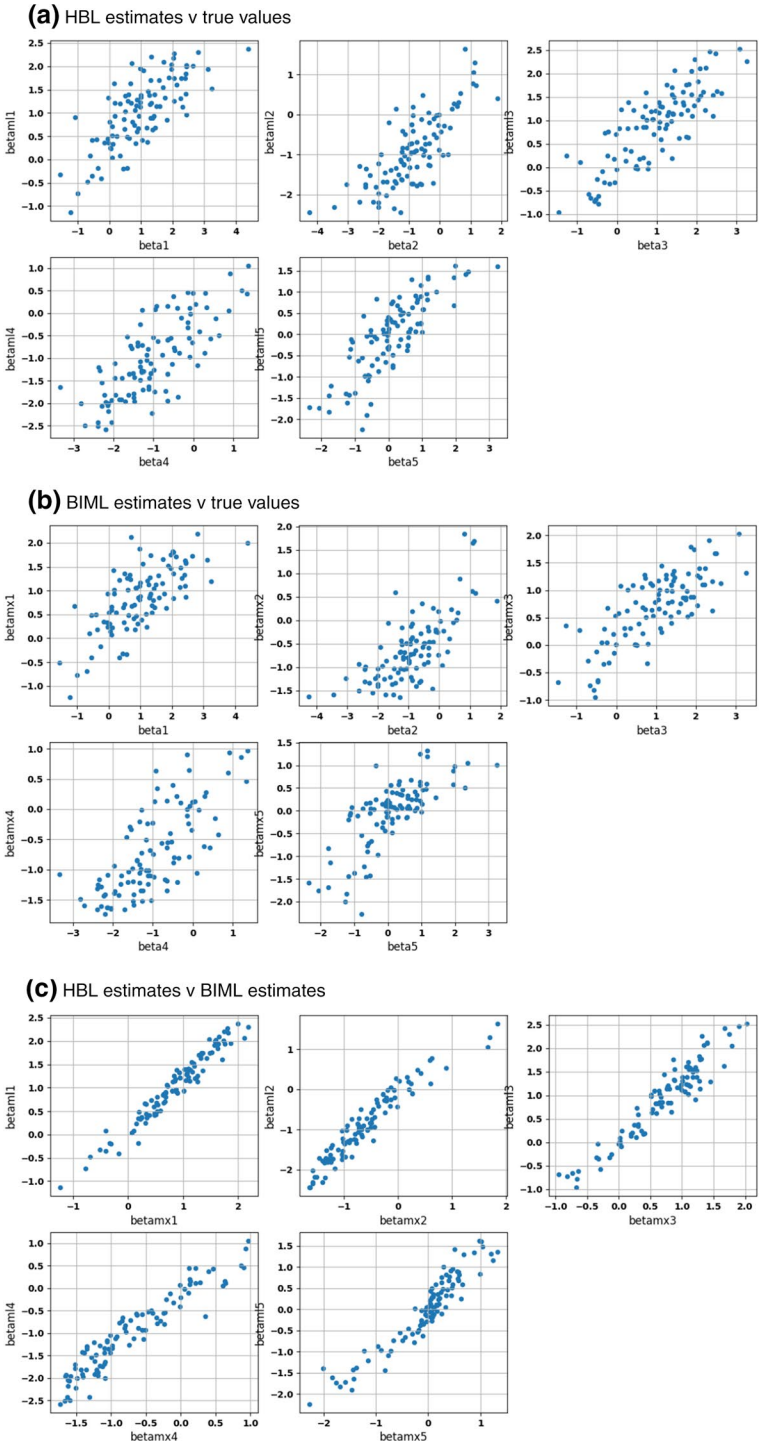

Fig. 5 Normally distributed parameters

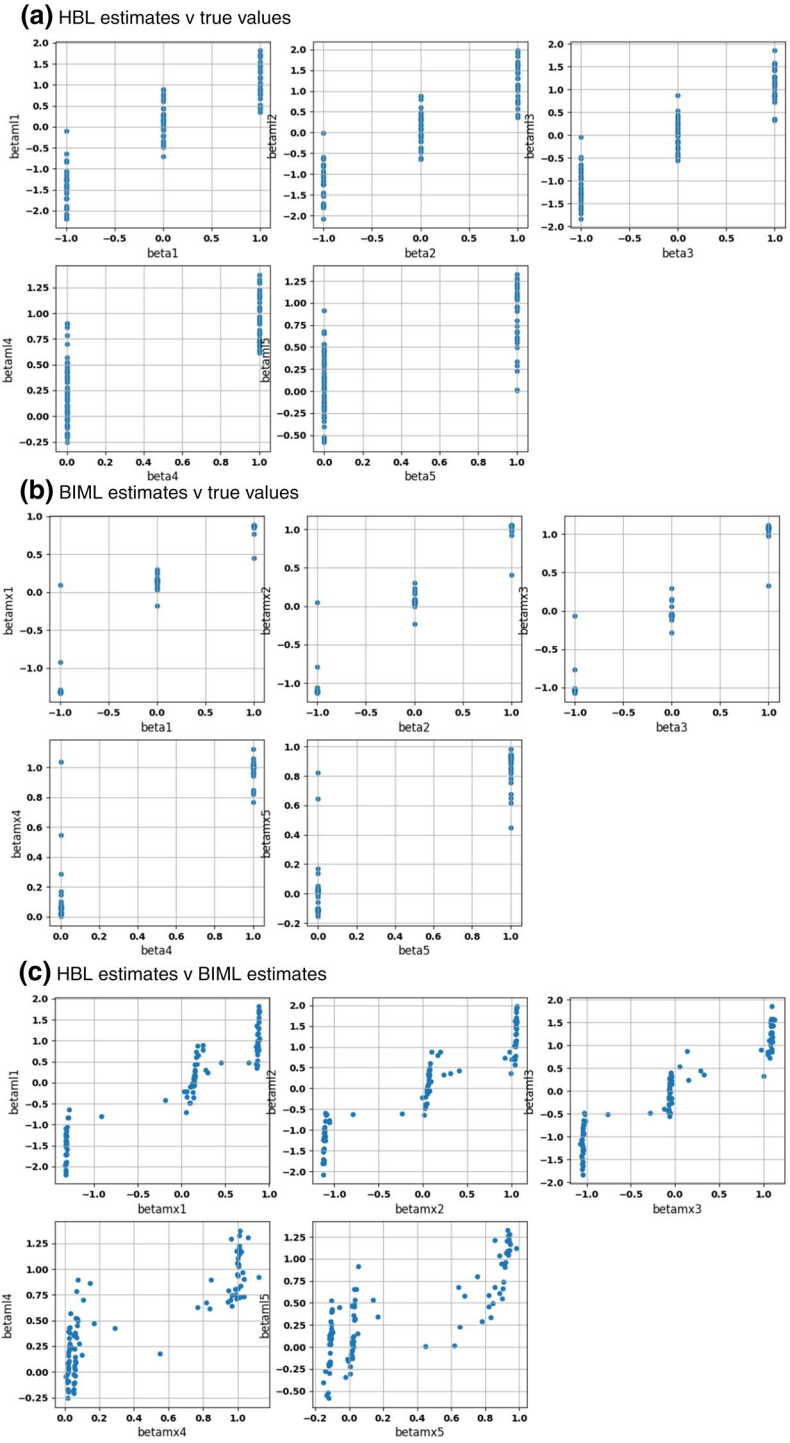

Fig. 6 Three mixtures

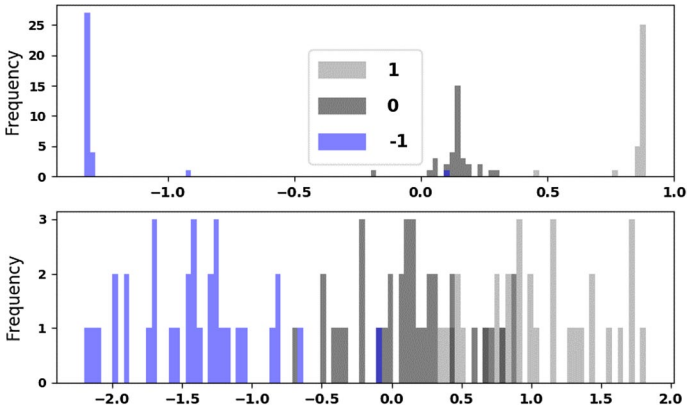

Fig. 7 HBL + BIML for coefficient 1, under three levels of utility

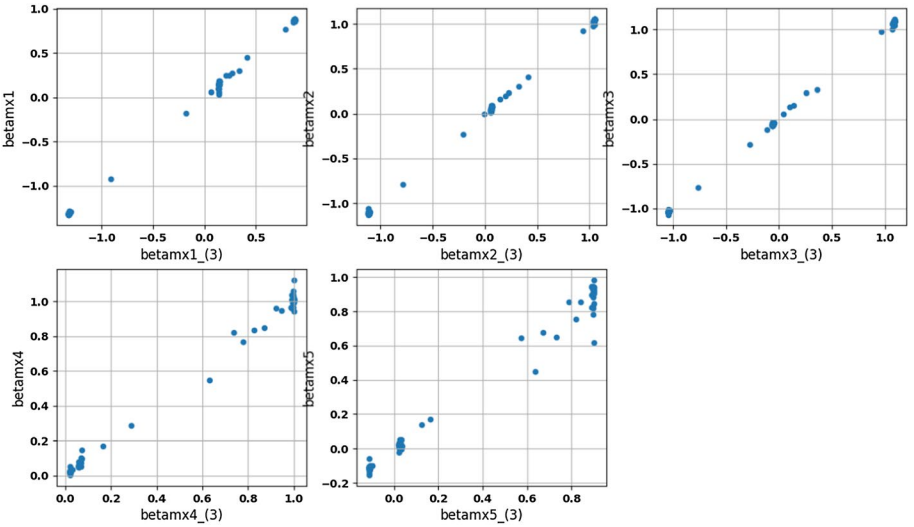

Fig. 8 BIML versus finite mixture with three mixtures

We now turn to the performance of the HBL and BIML for the second DGP outlined in Case 2 of Sect. 1.2. Our results are shown in Figs. 6 and 7. The same labelling structure is adopted as in Fig. 5.

In the same manner as for Fig. 5, in Fig. 6a the HBL estimates of the individual utility values are compared against the true (as specified above for Case 2) values for each of the five parameters. As is evident from Fig. 6a, the HBL is able to approximately separate each of the groups, but the distributions tend to be quite dispersed around the true value. The distributions also overlap to some extent. This is illustrated by the bottom histogram of Fig. 7 which has an almost uniform distribution of MUs. The bottom histogram in Fig. 7 corresponds to the upper left hand panel of Fig. 6a We only give a histogram for this first parameter since it is broadly indicative for all the other parameters. For those familiar with the HBL or classical variant of this model, these results are entirely predictable. The prior

assumption of a normal distribution for the parameters around a mean and given variance heavily pushes individual estimates to be generated in a manner consistent with this assumption, although with more observations for each individual the individual estimates would become more clustered around their true values (the prior normality assumption does not imply normally distributed posteriors for the individual MUs).

Next we turn to the performance of the BIML compared to the true values as represented by Fig. 6b and the bottom part of Fig. 7. As can be seen by Fig. 6b, there is not a perfect classification using the BIML, however, the estimates are much more densely packed around their true values than for the case of the HBL. This is to some extent obscured by the fact that the dots often represent the values for multiple respondents. The case of the first parameter is illustrated in the histogram at the bottom of Fig. 7, which corresponds the upper left and pain of Fig. 6b. We can see that there are three distinct densities close to the true values, with a "misclassification" of 2 out of 100 respondent utilities.

Finally, in Fig. 6c, we compare the HBL estimates and BIML estimates which further confirms the clustered nature of the BIML estimates (horizontal axis) compared to the HBL estimates. However, it also illustrates the same pattern as Case 1 where there is a tendency towards a monotonic relationship between the estimates of the two models.

Finally, turning to Fig. 8, we explore the consequences of specifying three mixtures only within the BIML. Within the Bayesian context this can easily be accomplished by accepting only those draws corresponding to have three mixtures only. This can be achieved providing this number has been visited enough times by the sampler, which is no problem for the current example. This also underlines the flexibility of the BIML approach, since researchers who wish to select a finite mixture can do so after estimation using the modal NOM. Note, this not an approach we recommend.

Figure 8 shows that there is the expected reduction in the dispersion of MU estimates at the values 0, − 1, and 1, but this is not a substantive reduction. This is consistent with the unrestricted BIML sampler proposing new mixtures periodically with higher dimensions that are populated by only a few (or only one) member, and being eliminated in subsequent draws. In short, their has only been a small reduction in selecting the correct number of classes by not setting the number of classes to its known value.

Summary

We have presented a comparative examination of the HBL and BIML models under two polar DGPs, one of which is completely consistent to the HBL and the other to a finite mixture model. In this study the BIML was able to handle both cases, not just the finite mixture case for which you would expect it to perform well. This highlights the potential of this model to flexibly estimate the marginal utilities within DCEs. It is not the purpose here to suggest based on a limited Monte Carlo experiment to claim that the BIML is superior. Clearly in many circumstances the HBL will perform as well, or better. However, the class of Dirichlet process models that underpin the BIML approach have had considerable acceptance within the statistics literature, yet they have not yet made their way onto the radar of most applied choice modellers. We believe that at the very least this class of models should be considered by applied choice modellers in the future.

References

Adekola O, Mitchell G, Grainger A (2015) Inequality and ecosystem services: the value and social distribution of Niger Delta wetland services. Ecosyst Serv 12:42–54

Agbagwa IO, Ndukwu BC (2014) Oil and gas pipeline construction-induced forest fragmentation and biodiversity loss in the Niger Delta, Nigeria. Nat Resour 5:698–718

Akpoghomeh OS, Badejo D (2006) Petroleum product scarcity: a review of the supply and distribution of petroleum products in Nigeria. OPEC Rev 30:27–40

Anejionu OCD, Ahiarammunnah P, Nri-ezedi CJ (2015) Hydrocarbon pollution in the Niger Delta: geographies of impacts and appraisal lapses in extant legal framework. Resour Policy 45:65–77

Anifowose B, Lawler DM, Horst D et al (2012) Attacks on oil transport pipelines in Nigeria: a quantitative exploration and possible explanation of observed patterns. Appl Geogr 32:636–651

Balcombe KG, Fraser IM, Lowe B et al (2016) Information customization and food choice. Am J Agric Econ 98:54–73

Bansal P, Daziano R, Achtnicht M (2018) Extending the logit-mixed logit model for a combination of random and fixed parameters. J Choice Model 27:88–96

Bauwens L, Carpantier J, Dufays A (2015) Autoregressive moving average infinite hidden Markov-Switching models. J Bus Econ Stat 35:162–182

Bazzani C, Palma MA, Nayga RM (2018) On the use of flexible mixing dis-distributions in WTP space: an induced value choice experiment. Aust J Agric Resour Econ 62:185–198

Bhat C, Lavieri P (2018) A new mixed MNP model accommodating a variety of dependent non-normal coefficient distributions. Theory Decis 84:239–275

BP (2015) BP Statistical Review of World Energy June 2014, 63rd edn. http://www.bp.com. Accessed 18 Sept 2018

Burda M, Harding M, Hausman J (2008) A bayesian mixed logit-probit model for multinomial choice. J Econom 147:232–246

Caputo V, Scarpa R, Nayga RM et al (2018) Are preferences for food quality attributes really normally distributed? An analysis using flexible mixing distributions. J Choice Model 28:10–27

Colombo S, Hanley N, Calatrava-Requena J (2005) Designing policy for reducing the off-farm effects for soil erosion using choice experiments. J Agric Econ 56:81–95

Daziano RA (2013) Conditional-logit bayes estimators for consumer valuation of electric vehicle driving range. Resur Energy Econ 35:429–450

Escobar MD, West M (1995) Bayesian density estimation and inference using mixtures. J Am Stat Assoc 90:557–558

Fiebig DG, Keane MP, Louviere J et al (2010) The generalized multinomial logit model: accounting for scale and coefficient heterogeneity. Market Sci 29:393–421

Flores NE (2002) Non-paternalistic altruism and welfare economics. J Public Econ 83:293–305

Frigyik, BA, Kapila A, Gupta MR (2010) Introduction to the Dirichlet distribution and related processes. UWEE Technical Report Number UWEETR-2010-0006, December 2010

Gelman A, Rubin DB (1992) Inference from iterative simulation using multiple sequences. Stat Sci 7(4):457–472

Greene WH, Hensher DA (2003) A latent class model for discrete choice analysis: contrasts with mixed logit. Transport Res B Methods 37:681–698

Haab TC, Interis MG, Petrolia DR et al (2013) From hopeless to curious? Thoughts on Hausman's 'Dubious to hopeless' critique of contingent valuation. Appl Econ Perspect P 35:593–612

Hensher DA, Greene WH (2003) The mixed logit model: the state of practice. Transportation 30:133–176

Idemudia U (2010) Corporate social responsibility and the Rentier Nigerian State: rethinking the role of government and the possibility of corporate social development in the Niger Delta. Can J Dev Stud 30:131–151

Idemudia U (2014) Corporate-community engagement strategies in the Niger Delta: some critical reflections. Extr Ind Soc 1:154–162

Imobighe MD (2011) Paradox of oil wealth in the Niger-Delta region of Nigeria: how sustainable is it for national development? J Sustain Dev 4:16–168

Jacobsson F, Johannesson M, Borgquist L (2007) Is altruism paternalistic? Econ J 117:761–781

Jin X, Maheu JM (2016) Bayesian semiparametric modeling of realized covariance matrices. J Econom 192:19–39

Kosenius A, Ollikainen M (2013) Valuation of environmental and societal trade-offs of renewable energy sources. Energy Policy 62:1148–1156

Longo A, Markandya A, Petrucci M (2008) The internalization of externalities in the production of electricity: willingness to pay for the attributes of a policy for renewable energy. Ecol Econ 67:140–152

McAuliffe JD, Blei DM, Jordan MI (2006) Nonparametric empirical Bayes for the Dirichlet process mixture model. Stat Comput 16:5–14
McConnell K (1998) Does altruism undermine existence value? J Environ Econ Manag 32:22–37
Neal RM (2000) Markov chain sampling methods for Dirichlet process mixture models. J Comput Graph Stat 9:249–265
Nigerian Bureau of Statistics (NBS) (2012) Annual Abstract of Statistics. Federal Republic of Nigeria, NBS. http://www.nigerianstat.gov.ng. Accessed 18 Sept 2018
Nigerian Bureau of Statistics (NBS) (2015) Unemployment/Under-Employment Watch Q1 2015. http://www.nigerianstat.gov.ng. Accessed 18 Sept 2018
Omiyi B (2001) Shell Nigeria corporate strategy for ending gas flaring. In: Proceedings of the paper presented at a seminar on gas flaring and poverty alleviation, June 18–19, Oslo. Shell Petroleum Development Company of Nigeria Ltd, pp 2–13
Omofonmwan SI, Odia LO (2009) Oil exploitation and conflict in the Niger-Delta region of Nigeria. Hum Ecol 26:25–30
Ritchie J, Lewis J (2011) Qualitative research practice, a guide for social science students and researchers. Sage publications, London
Scarpa R, Rose JM (2008) Design efficiency for non-market valuation with choice modelling: how to measure it, what to report and why. Aust J Agric Resour Econ 52:253–282
Shell Petroleum Development Company (2012) Oil spills in the Niger Delta-monthly data for 2011. http://www.shell.com.ng/environment-society/environment-tpkg/oil-spills/data-2011.html. Accessed 18 Sept 2018
Songi O (2015) Defining a path for benefit sharing arrangements for local communities in resource development in Nigeria: the Foundations, Trusts and Funds (FTFs). Model J Energy Nat Resour Law 33:147–170
Train KE (2003) Discrete choice methods with simulation. Cambridge University Press, New York
Train KE (2008) EM algorithms for nonparametric estimation of mixing distributions. J Choice Model 1:40–69
Train KE (2016) Mixed logit with a flexible mixing distribution. J Choice Model 19:40–53
United Nations Development Report (UNDP) (2006) Niger-Delta Development Human Report
United Nations Development Program (UNDP) (2013) Nigeria's 2013 Millennium Development Goals Report. www.mdgs.gov.ng. Accessed 18 Sept 2018
United Nations Development Program (UNDP) (2015) Millennium Development Goals. End-Point Report 2015, Nigeria. www.mdgs.gov.ng. Accessed 18 Sept 2018
Urama KC, Hodge I (2006) Participatory environmental education and willingness to pay for river basin management: empirical evidence from Nigeria. Land Econ 82:542–561
Walker JL, Ben-Akiva M, Bolduc D (2007) Identification of parameters in normal error component logit-mixture (NECLM) models. J Appl Econ 22:1095–1125
Zhang J, Adamowicz W, Dupont DP et al (2013) Assessing the extent of altruism in the valuation of community drinking water quality improvements. Water Reour Res 49:6286–6297

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.